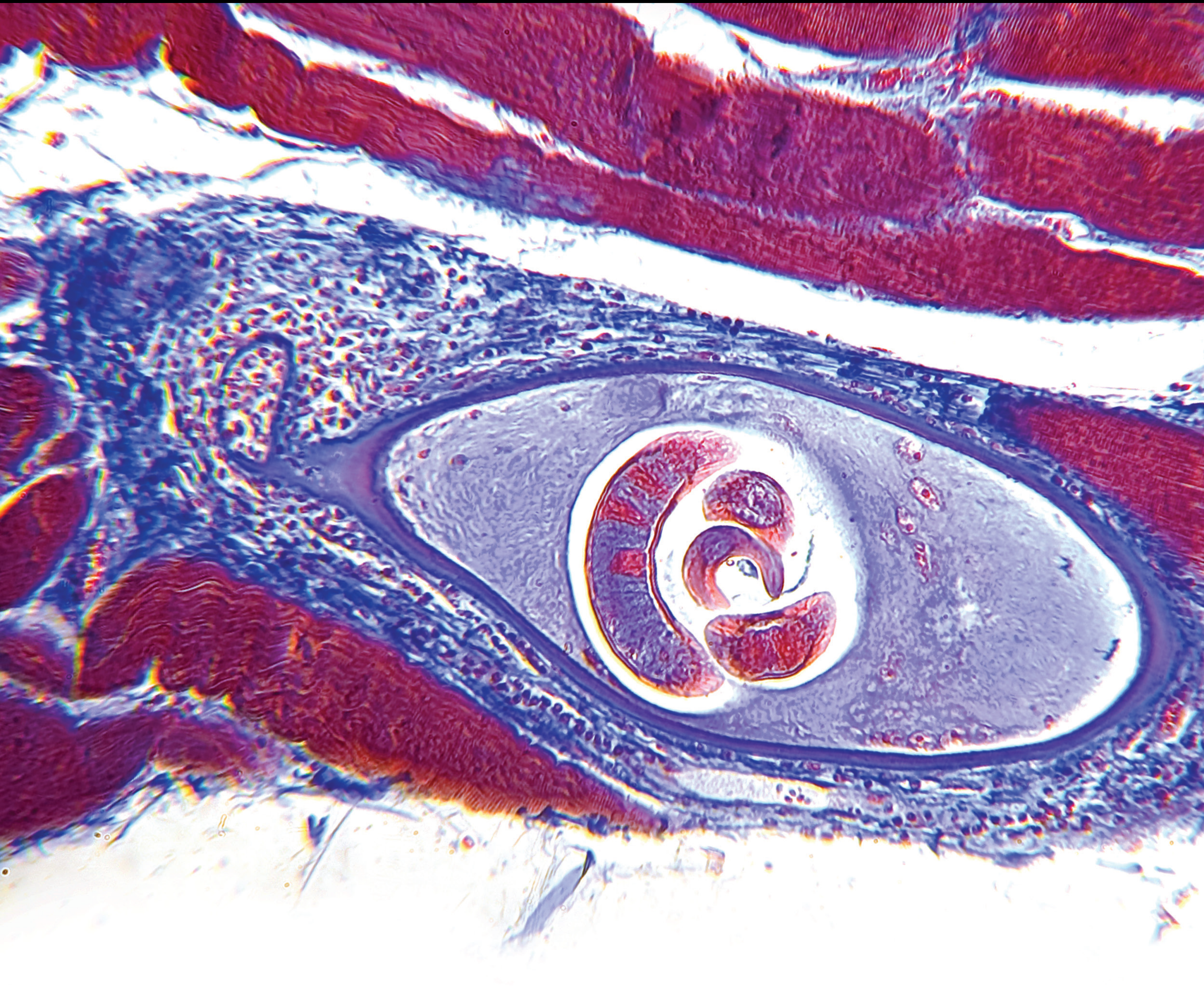
Gastroenterology Research and Practice

Artificial Intelligence in Gastroenterology

Lead Guest Editor: Mohit Girotra
Guest Editors: Andrew Ofosu, James H. Tabibian, Monique Barakat, and Soumya Jagannath Mahapatra



Artificial Intelligence in Gastroenterology

Artificial Intelligence in Gastroenterology

Lead Guest Editor: Mohit Girotra
Guest Editors: Andrew Ofosu, James H. Tabibian, Monique Barakat, and Soumya Jagannath Mahapatra

Contents

Hindawi
Gastroenterology Research and Practice
Volume 2023, Article ID 3228832, 7 pages
https://doi.org/10.1155/2023/3228832

Review Article

Artificial Intelligence in Inflammatory Bowel Disease Endoscopy: Advanced Development and New Horizons

Yu Chang, Zhi Wang, Hai-Bo Sun, Yu-Qin Li, and Tong-Yu Tang

Department of Gastroenterology, First Hospital of Jilin University, Changchun, 130000 Jilin, China

Correspondence should be addressed to Yu-Qin Li; liyuq@jlu.edu.cn and Tong-Yu Tang; tangty@jlu.edu.cn

Received 2 November 2022; Revised 28 November 2022; Accepted 2 December 2022; Published 17 April 2023

Academic Editor: James H. Tabibian

Inflammatory bowel disease (IBD) is a complex chronic immune disease with two subtypes: Crohn's disease and ulcerative colitis. Considering the differences in pathogenesis, etiology, clinical presentation, and response to therapy among patients, gastroenterologists mainly rely on endoscopy to diagnose and treat IBD during clinical practice. However, as exemplified by the increasingly comprehensive ulcerative colitis endoscopic scoring system, the endoscopic diagnosis, evaluation, and treatment of IBD still rely on the subjective manipulation and judgment of endoscopists. In recent years, the use of artificial intelligence (AI) has grown substantially in various medical fields, and an increasing number of studies have investigated the use of this emerging technology in the field of gastroenterology. Clinical applications of AI have focused on IBD pathogenesis, etiology, diagnosis, and patient prognosis. Large-scale datasets offer tremendous utility in the development of novel tools to address the unmet clinical and practice needs for treating patients with IBD. However, significant differences among AI methodologies, datasets, and clinical findings limit the incorporation of AI technology into clinical practice. In this review, we discuss practical AI applications in the diagnosis of IBD via gastroenteroscopy and speculate regarding a future in which AI technology provides value for the diagnosis and treatment of IBD patients.

1. Introduction

Artificial intelligence (AI) represents the capacity of machines to imitate human intelligence. Major aspects of AI applications in medicine include computational intelligence, gene sequencing, intelligent diagnosis, and medical robotics. Currently, the application of AI technology to gastrointestinal endoscopy is increasing rapidly. Compared with professional endoscopists, AI technology has been found to have superior accuracy for analyzing and processing large volumes of medical data. Machine learning (ML), which is essential in the implementation of AI, is the process of using algorithms to guide a computer to use known data to obtain an appropriate model. This model can then be used to assess new situations. As a branch of contemporary statistics, ML is particularly useful for analyzing complex data [1]. There are four types of ML: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. In supervised learning, an algorithm is given a dataset that includes questions and correct answers, and the machine learns how to predict the correct answers to future questions by analyzing the data. Deep learning (DL) is a subset of ML [2] that has received attention in the field of medical imaging science as a fully automated, fast, and accurate imaging analysis solution. A typical example of DL is the convolutional neural network (CNN), which can be used to address image-based problems in medicine. Given their great utility, CNNs have become widely applied in medical imaging [2].

Inflammatory bowel disease (IBD) is a chronic complex inflammatory disease with increasing incidence globally. IBD prognosis is closely related to the healthcare system [3]. The two subtypes of IBD, ulcerative colitis (UC) and Crohn's disease (CD), are typical complex diseases characterized by chronic and heterogeneous presentations. They are induced by interactions between genomic, environmental, microbial, and immunological factors [4]. The accurate diagnosis of IBD has long been a challenge for gastroenterologists. However, new IBD diagnostic techniques include combinations of methods such as gastrointestinal endoscopy, molecular pathology, epigenetics, metabolomics, and proteomics [5]. The state-of-the-art endoscopic imaging techniques and novel biomarkers provide new approaches to the differential diagnosis of IBD. Over the past few years, the importance of endoscopy in the diagnosis, treatment, and monitoring of IBD has been established. For example, dye-chromo endoscopy (DCE) and virtual chromo endoscopy (VCE) are often used in the endoscopic surveillance of IBD [6]. In DCE, topical dye is sprayed on the colon wall, enhancing the visualization of mucosal morphology. This technique is now recognized as the gold standard for diagnosing hyperplasia and has been included in the recommendations of international diagnosis guidelines. VCE, which is also included in international guidelines, involves the digitization of endoscopic images, enabling tissue surface details to be enhanced with high accuracy. Accordingly, it can function as an alternative to DCE. Despite the utility of these techniques, differences in the specific methods used as well as the quality and subsequent interpretation of diagnostic components can significantly affect the diagnosis and treatment outcomes of IBD among gastroenterologists.

Today, progress in AI technology has dramatically enhanced the ability of clinicians and researchers to analyze, manipulate, interpret, and apply large data sets. The amount of data from clinical trials, medical imaging, and genetic research (genomic, transcriptomic, and proteomic) is rapidly increasing [7]. Without appropriate methods for interpretation, it is difficult to combine large amounts of clinical data with genetic data for detailed analysis in clinical practice. AI and ML can be used to quickly analyze these datasets, enabling clinicians to implement stratified management of patients in terms of risk assessment, diagnosis, treatment, and prognosis. Thus, AI and ML have enabled more accurate and standardized endoscopic treatment measures. With the enhanced application of AI in IBD treatment and diagnosis, endoscopic procedures have become increasingly specialized and currently include the evaluation of endoscopic disease activity, monitoring of cancerous lesions, and capsule endoscopy (CE) for the diagnosis of CD [8]. The purpose of this review is to summarize the current application of AI techniques to the diagnosis and treatment response prediction of IBD and to discuss future directions in the applications of AI to IBD endoscopy.

2. Classification of AI

As a popular research topic, interest in AI has grown rapidly in the medical community over recent years. Accordingly, the application of AI to IBD treatment has led to significant progress in computer-assisted diagnosis and therapy [9]. Many computer algorithms have been developed to assist in gastroscopy. ML, as a form of AI, can be used to facilitate algorithmic self-improvement based on experience and without human supervision. Specifically, the ML algorithm learns from inputted data sets, and identifies behavior patterns or generates predictive models [10]. The application of ML in endoscopic IBD monitoring can be realized by analyzing still images. DL, as a subset of ML, can be used to handle complicated learning algorithms. CNNs, as a type of DL, are becoming the leading technology for image processing.

2.1. Machine Learning. In the near future, ML-based models are expected to be employed by a large number of clinicians for image recognition and analysis. In the field of IBD research, ML has been used to determine whether the mucosa is healing in patients with UC [11] and to classify subtypes of pediatric colonic IBD [12]. Huang et al. invented a computer-aided diagnostic system based on ML and DL (DLML-CAD). The concept underlying the system is related to transfer learning, in which a classifier is trained to extract the desired features of images using a network that has been pre-trained using millions of non-medical images. The investigators chose the deep neural networks(DNN), support vector machine, and k-nearest neighbor network models as classifiers for the DLML-CAD, and classified hundreds of images as Mayo endoscopic subscore (MES) 0–1 or MES 2–3. The DLML-CAD has reached or even slightly surpassed the diagnostic level of IBD endoscopists. The system can identify and analyze colon endoscopic images to accurately determine the degree of mucosal healing (ML). It can also be used to evaluate the mucosa in different areas of the colon during colonoscopies in patients with UC [11]. Dhaliwal et al. collected clinical, endoscopic, radiographic, and histological data from 74 patients with colonic IBD, trained a random forest classifier on the complete dataset, and used ML to identify three histological features and four endoscopic examinations that could be used to distinguish colonic UC from CD [12]. According to the existing data, the ML model performs very well in the field of IBD diagnostics and treatment. However, its widespread application in clinical settings is still uncertain as it is still in the clinical trial stage.

2.2. Deep Learning. Given the rigorous training required to accurately assess endoscopic inflammation and the limited number of endoscopists with specialized training, DL may offer many advantages in the clinical evaluation of IBD patients. According to a relevant study, small bowel ulcers in patients with CD, as revealed via video CE, can be effectively monitored by DL models when the area under the receiver operating characteristic curve is between 0.94 and 0.99 [13]. Takenaka et al. also validated the use of DL algorithms for assessing disease activity using a UC endoscopic severity score [14], and DL algorithms appear to be of trans-generational significance for the rapid acquisition and analysis of images during endoscopy.

2.3. Convolutional Neural Networks. CNNs are a type of DL algorithm that have an enormous impact on the field of computer vision, with image analysis accuracy comparable to that of professional physicians. CNNs simulate neuronal networks in the brain by combining many data inputs, weights, and biases. These systems are extremely useful in detecting intestinal ulcers, erosions, and strictures [15, 16]. Ding et al. validated the ability of a CNN-based ML model to efficiently identify and analyze small bowel capsule endoscopic images (SB-CE) and concluded that it could be a powerful tool to help experienced endoscopists quickly and accurately analyze and classify small bowel lesions [17], whereas Klang et al. monitored endoscopic ulcers in video capsule endoscopic images and successfully identified CD patients via an AI system based on CNNs [13]. However, no algorithms have performed in a superior manner to professional endoscopists. Furthermore, AI algorithms for the long-term monitoring of IBD patients have not yet been developed. Many such models are in the clinical trial stage and, thus, are not widely used in large-scale clinical practice.

3. AI in the Diagnosis of IBD

The diagnosis of IBD is a highly sophisticated process, because it must be carefully discriminated from other diseases. Clinicians must consider the patient's history and clinical presentation, and conduct a series of endoscopic and histological examinations. Endoscopic techniques are the cornerstone of IBD diagnosis, and endoscopists use a variety of methods to classify and analyze mucosal erosions, inflammation, and ulcers. These include histological examinations such as microscopic biopsy combined with a series of imaging examinations such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). The endoscopic diagnosis of IBD requires highly specialized operators with extensive training and experience. However, inter-operator variability in endoscopic evaluation is inevitable. Such variation may be addressed by AI-based computer-assisted systems.

3.1. AI in Diagnosing UC. As endoscopic remission is a therapeutic goal for UC patients, endoscopy has been widely used to assess the activity and efficacy of treatments [18]. Histological remission has also become an increasingly important therapeutic goal for UC [19]. However, to determine the status of histological remission targets, pathologists must evaluate mucosal inflammation. This can result in observer variability and, thus, data heterogeneity. Computer-aided systems can help physicians to determine the degree of inflammation and ML during UC endoscopy, resulting in a more accurate assessment of histological remission.

One study demonstrated that the predictions of a model based on DL algorithms regarding the severity of UC during endoscopy were largely consistent with those of experienced human evaluators [20]. Thus, such algorithms may improve the assessment and treatment of UC according to endoscopic data. Sutton et al. investigated the use of DL algorithms in differentiating UC from other intestinal diseases and evaluating the severity of UC endoscopic ulcers [21]. They used a dataset containing 851 images of UC patients that had been labeled and graded by professional endoscopists using the MES. Other studies on the automatic grading of UC endoscopic images have also shown advantages of DL models. For instance, Takenaka et al. evaluated endoscopic images of UC patients using a model based on a DL network [14]. Their model was highly accurate (90.1%) when evaluating endoscopic images of 40,758 UC patients with endoscopic remission who had received a UC Endoscopic Severity Index Score (UCEIS) of 0. Similarly, Yao et al. modified the traditional full-motion video operation model by segmenting the endoscopic video into 1-frame-per-second image stacks and then automatically rotating, fragmenting, and preprocessing the images to conform to a standard scale [22]. The resulting model was able to automatically generate MESs for patients. Gottlieb et al. utilized CNNs to collate 795 full-length endoscopic videos, and collected and cleaned the data so that the endoscopic Mayo score and the UCEIS scores could be used for prediction and analysis [23]. To reduce the subjectivity of operator-assessed endoscopic image activity scoring, Bossuyt et al. developed a non-operator-dependent objective endoscopic scoring system to assess UC disease activity. The system was based on the red, green, and blue pixel values and endoscopic image recognition, and it performed UC activity assessment with high operability and objectivity [24]. More details can be shown clearly in Table 1.

The DNN model for UC assessment can score endoscopic images with very high accuracy. Indeed, many studies have shown that DNNs can assess the activity and remission of intestinal mucosal inflammation through endoscopic images alone, eliminating the need for biopsies and reducing the need for pathology. The objectivity and consistency of such systems are comparable to that of professional endoscopists. However, many challenges must be addressed before AI can be formally applied to large-scale clinical practice. For example, the data used to train DNNs are defined, organized, and structured by humans based on collected images, which can lead to problems such as the human assessment of discrepancies.

3.2. AI in Diagnosing CD. Because the intestinal inflammatory permeation in CD includes the small intestine, especially the terminal ileum, CE is often used to detect CD when the applications of conventional colonoscopy are limited. However, CE can generate videos that are 8–10 hours in length, which makes frame-by-frame inspection and analysis very time-consuming and labor-intensive for endoscopists [17]. In response to this, a number of CNN algorithms have been developed.

In 2019, Ding et al. validated the ability of CNN-based algorithms to analyze SB-CE images with greater accuracy and sensitivity than conventional methods and reported a significantly reduced reading time compared with routine analysis by gastroenterologists [17]. Indeed, this technique has substantial benefits for gastroenterologists in terms of efficiency and convenience in organizing and analyzing image information. In 2019, Aoki et al. developed a CNN-based algorithm that could automatically detect erosions and ulcers from capsule endoscopic images. They tested their model using 10,440 test images and found the sensitivity, specificity, and accuracy to be 88.2%, 90.9%, and 90.8%, respectively [25]. In 2019, Klang et al. designed and trained a CNN to randomly segment 17,640 capsule endoscopic CD images [13]. They achieved good results with an area under the curve of 0.99 and an accuracy of 95.4–96.7%. In 2021, Klang et al. further evaluated a CNN-based AI system. They explored the accuracy of the system in comparing CD stenosis and different degrees of ulcers and found that such systems may be capable of automatic identification and

Table 1: Endoscopic AI applications in diagnosing UC.

Author	Year	Data source	Purpose	Results
Sutton et al.	2022	851 images of UC patients	To differentiate UC from other intestinal diseases and to evaluate the severity of UC endoscopic ulcers	The accuracy (87.50%) and area under the curve (AUC, 0.90)
Takenaka et al.	2020	40,758 images of endoscopies and 6,885 biopsy outcomes of 2012 UC patients	To create a deep neural network system for analyzing endoscopic images of UC patients	The remission in endoscopy with 90.1% accuracy (95% and 89.2–90.9%)
Najarian et al.	2021	Video of the clinical trial set with 51 high resolution and 264 tests	To trial a fully automated video system for analyzing and grading endoscopic disease in UC	Automated Mayo endoscopic subscores (MES) scoring of clinical trial videos correctly differentiated between remission and active disease in 83.7% of cases
Gottlieb et al.	2021	795 full-length endoscopy videos of 249 patients	To verify a deep learning, algorithm can be trained to predict levels of UC severity from full-length endoscopy videos	(0.787–0.901) for endoscopic Mayo Score (eMS) and 0.855 (95% confidence interval, 0.80–0.91) for UCEIS
Bossuyt et al.	2020	29 consecutive patients with UC and 6 healthy controls	To develop an operator- independent computer-based tool to determine UC activity based on endoscopic images	RD correlated with Robarts histological index (RHI) ($r = 0.74$, $p < 0.0001$), MES ($r = 0.76$, $p < 0.0001$), and UC endoscopic index of severity scores ($r = 0.74$, $p < 0.0001$)

grading of CD in the near future [26]. In 2021, Barash et al. successfully developed a DL algorithm for CE that could automatically grade CD ulcers, demonstrating that CNN-assisted CE readings have high utility in the diagnoses and monitoring of CD patients [27]. More details can be shown clearly in Table 2.

Given the findings mentioned above, AI-based DL algorithms, especially CNNs, appear to strongly improve the accuracy of CE analysis while greatly reducing the time required for endoscopists to examine images and videos. However, these DL-based detection algorithms are performed at the level of individual images but not at the entire video level. This means that the samples for these experiments are from retrospective studies and do not fully resemble the performance of video CE. Future researchers working with AI technologies are likely to develop algorithms that can automatically evaluate CE videos, resulting in accurate scoring systems that can be widely used in clinical practice. As the technology continues to advance, the combination of AI and CE is expected to significantly impact endoscopy. In the future, AI-assisted CE may be able to perform rapid and systematic examinations of entire intestinal lesions in less than 30 minutes. In the meantime, CE systems that facilitate patient diagnosis, treatment, and biopsy appear to be driving a revolution in endoscopy technology.

3.3. AI in Cancer Surveillance of IBD. Given that IBD is a long-term chronic condition, patients have a greatly increased risk of colorectal cancer compared with the general population. One study revealed that the existence of low-grade dysplasia (LGD) in the intestine of IBD patients served as a high-risk factor for progressing to high-grade dysplasia (HGD) or even colorectal cancer. Therefore, once LGD is detected, patients should undergo careful endoscopic screening [28]. In IBD patients, endoscopic screening for colorectal cancer, LGD, and HGD is currently performed mainly via stained endoscopy plus endoscopic resection or biopsy [29]. To date, no AI systems have been developed for the long-term monitoring of patients with IBD colitis. In 2020, Maeda et al. reported the first case of AI-assisted detection of colitis-associated neoplasms [30]. Their patient was 72 years old and possessed an 18-year history of colitis. In 2021, Maeda et al. designed an automated AI algorithm-based polypectomy surveillance system for use during surveillance colonoscopy. Their system was able to clearly and efficiently identify colonic lesions in non-IBD patients, which confirmed the feasibility of the system for helping non-endoscopic specialists identify and detect long-term heterogeneous growths in IBD patients.

4. AI in the Treatment of IBD

Histological remission has gradually become a therapeutic goal for UC, replacing previous goals such as endoscopic and symptomatic remission. Accurate assessments of the degree of histologic remission and intestinal inflammatory activity may allow for more precise and efficient treatments as well as reduce the need for multiple repeated gastrointestinal examinations [9]. Unfortunately, clinical practice is currently lacking a simple and easy standard for evaluating histological remission. In 2022, Villanacci et al. developed a semi-supervised AI inductive transfer learning system consisting of two modules. Their goal was to apply the simplified neutrophil-only Paddington International virtual ChromoendoScopy ScOre histological remission index (PHRI) developed and validated by pathologists to a computer-aided diagnostic system. When comparing the evaluation results produced by the AI with those generated by pathologists, the AI model was found to be highly sensitive and specific in determining the presence of neutrophils, indicating that it could be an excellent support in determining whether patients had achieved histological remission [31]. Gui et al. also developed a CNN-based computer-aided UC histological diagnosis and scoring system for

Table 2: Endoscopic AI applications in diagnosing CD.

Author	Year	Data source	Purpose	Results
Ding et al.	2019	4,206 abnormalities in 3,280 patients	To analyze SB-CE images with greater accuracy and sensitivity than conventional methods	99.88% sensitivity in the per-patient analysis (95% confidence interval [CI], 99.67–99.96)
Aoki et al.	2019	113,426,569 images from 6,970 patients	To develop a CNN-based algorithm that could automatically detect erosions and ulcers from capsule endoscopic images	With 99.88% sensitivity in the per-patient analysis (95% CI, 99.67–99.96) and 99.90% sensitivity in the per-lesion analysis (95% CI, 99.74–99.97)
Klang et al.	2019	17,640 CE images from 49 patients	To evaluate a deep learning algorithm for the automated detection of small-bowel ulcers in Crohn's disease (CD) on capsule endoscopy (CE) images	AUCs (area under the curve of 0.99 and accuracies ranging from 95.4–96.7%
Klang et al.	2021	27,892 CE images	To prove the ability of deep neural networks to identify intestinal strictures on CE images of Crohn's disease (CD) patients	A differentiation between strictures and normal mucosa (area under the curve [AUC], 0.989)
Barash et al.	2021	17,640 CE images from 49 patients	To develop a deep learning algorithm for automated grading of CD ulcers on CE	The accuracy of the algorithm was 0.91 (95% CI) for distinction of 3 different grades

identifying disease activity in UC patients based on PHRI item scores. Their system was designed to not only avoid the subjectivity of pathologists in determining the degree of inflammatory activity but also to reduce the degree of difficulty of judgments regarding patient status. The PHRI is a simple and reproducible scoring system that is well suited for large-scale applications in clinical practice. It is effective in assessing UC endoscopic activity and also allows physicians to make more accurate conclusions about the status of histological remission [32]. However, although the studies reviewed above proposed using the PHRI score as a histological diagnostic and grading tool for UC, their work contains certain limitations. For example, in the study by Gui et al., the follow-up protocol excluded endoscopic and histological reassessment, the duration of the follow-up period was relatively short, and the investigators did not calculate overall PHRI scores for different regions of the intestine.

Given the complexity and chronicity of IBD as an immune intestinal disease, as well as the multiple factors that influence patient outcomes, the goal of many current therapies is symptomatic relief and ML. To this end, various targeted drugs and biological agents have been used in clinical practice. Maintaining the integrity of the intestinal epithelial barrier has become a research hotspot and an emerging therapeutic target in the biomedical field. Sahoo et al. used an ML approach to build a network that identified a pathway rich in gene clusters that maintain intestinal epithelial barrier integrity, leading to the identification of a top intestinal barrier protector for the treatment of IBD [33]. Experts in the field are expected to use similar methods to invent additional drugs and therapeutics that maintain or repair the intestinal epithelial barrier, and this is likely to become an exciting research area.

For many years, minimally invasive techniques have been used to treat IBD. Currently, robotic surgery is becoming an efficient and precise complement to, and potentially a future replacement for, minimally invasive surgery. There is growing evidence indicating that surgical robots have significant advantages over minimally invasive techniques such as laparoscopic surgery, including better patient safety, reduced surgical complications, and shortened prognosis [34]. In 2020, Hota et al. investigated the differences in perioperative and treatment outcomes between open, laparoscopic, and robotic surgeries for treating CD [35]. They selected a database containing data from 5,158 patients with CD, utilized Convolutional point transformer (CPT) codes to determine the procedures used for patient ileal resection, compared the incidence of anastomotic fistula between the three surgical approaches, and applied multivariate analysis to derive a 95% confidence interval for the dominance ratio. They found that robotic surgery was a non-inferior treatment for both colonic resection in UC recipients and ileostomy in CD patients [34, 35].

5. Risks and New Horizons

AI applications focused on disease prediction and cancer surveillance, and the diagnosis and treatment of IBD have been found to be extremely reliable and efficient. In the future, AI technologies may be able to completely replace endoscopists for decision-making and treatment, or alternatively, endoscopists may act as assistants to AI systems. For this to occur, prerequisites for the large-scale use of AI algorithms in clinical practice must be met, and ethical guidelines regarding patient safety must be put in place. Endoscopy clinics and academies will also need to develop emergency measures for malfunctions during AI treatment and remedies for treatment errors. In addition to the above, the biases from developers of AI-based algorithms must be considered, as most datasets are human-trained. Finally, the impact of different dataset types and analyses must be compared to accurately predict the status and value of AI algorithms in IBD clinical practice.

These above-mentioned risk factors should not prevent us from continuing to research AI algorithms, refine system functionality, and work to realize the full potential of AI-assisted medicine. The positive impact of AI and DL in gastroenterology is substantial, and many publicly available datasets are available for further comparison and analysis by researchers. In terms of future development, new AI algorithms for long-term monitoring of colorectal cancer in IBD patients are urgently needed to improve the prediction of cancer risk and time-course of treatment. Furthermore, randomized controlled trials are necessary to investigate the benefits and feasibility of using AI in the clinical management of IBD patients compared with general clinical management, especially in terms of variations in treatment measures, outcomes, and treatment costs. Currently, endoscopists are collaborating with algorithm developers with the goal of using large data sets for medical imaging AI. Indeed, the creation of data sets for collecting and training AI to capture new types of images will require the expertise and experience of endoscopists.

6. Conclusions

This review summarizes recent applications of AI in the endoscopic examination and treatment of IBD and predicts possible future directions for the use of AI in treating patients with IBD. We expect that AI will soon become a cornerstone of endoscopy and IBD treatment. Besides, it will be necessary to translate the large amount of current exploratory data into evidence that can be applied to clinical practice before AI could be widely used in clinical settings. This will require not only the rapid development of AI but also the cooperation and commitment of endoscopists, specialists, and societies.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Yu Chang and Zhi Wang contributed equally to this work. Tong-Yu Tang and Yu-Qin Li guided the conception of this review, and Hai-Bo Sun inspected the errors in the manuscript and fixed them.

Acknowledgments

The present study was supported by Department of Science and Technology of Jilin Province Agency (grant no. 20210402013GH) and Department of Finance of Jilin Province (grant no. JLSCZD2019-035).

References

[1] E. Mossotto, J. J. Ashton, T. Coelho, R. M. Beattie, B. D. MacArthur, and S. Ennis, "Classification of paediatric inflammatory bowel disease using machine learning," *Scientific Reports*, vol. 7, no. 1, p. 2427, 2017.

[2] J. Choi, K. Shin, J. Jung et al., "Convolutional neural network technology in endoscopic imaging: artificial intelligence for endoscopy," *Clinical Endoscopy*, vol. 53, no. 2, pp. 117–126, 2020.

[3] S. C. Ng, H. Y. Shi, N. Hamidi et al., "Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies," *Lancet*, vol. 390, no. 10114, pp. 2769–2778, 2017.

[4] H. S. P. de Souza, C. Fiocchi, and D. Iliopoulos, "The IBD interactome: an integrated view of aetiology, pathogenesis and therapy," *Nature Reviews. Gastroenterology and Hepatology*, vol. 14, no. 12, pp. 739–749, 2017.

[5] G. E. Tontini, M. Vecchi, L. Pastorelli, M. F. Neurath, and H. Neumann, "Differential diagnosis in inflammatory bowel disease colitis: state of the art and future perspectives," *World Journal of Gastroenterology*, vol. 21, no. 1, pp. 21–46, 2015.

[6] A. Dal Buono, R. Gabbiadini, F. Furfaro, M. Argollo, T. V. T. Trigo, and A. Repici, "Endoscopic surveillance in inflammatory bowel diseases: selecting a suitable technology," *Frontiers in Medicine*, vol. 9, p. 855652, 2022.

[7] I. S. Stafford, M. Kellermann, E. Mossotto, R. M. Beattie, B. D. MacArthur, and S. Ennis, "A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases," *NPJ Digital Medicine*, vol. 3, p. 30, 2020.

[8] G. E. Tontini, A. Rimondi, M. Vernero, H. Neumann, M. Vecchi, and C. Bezzio, "Artificial intelligence in gastrointestinal endoscopy for inflammatory bowel disease: a systematic review and new horizons," *Therapeutic Advances in Gastroenterology*, vol. 14, p. 17562848211017730, 2021.

[9] G. E. Tontini and H. Neumann, "Artificial intelligence: thinking outside the box," *Best Practice and Research Clinical Gastroenterology*, vol. 52–53, p. 101720, 2021.

[10] G. Chartrand, P. M. Cheng, E. Vorontsov et al., "Deep learning: a primer for radiologists," *Radiographics*, vol. 37, no. 7, pp. 2113–2131, 2017.

[11] T. Y. Huang, S. Q. Zhan, P. J. Chen, C. W. Yang, and H. H. Lu, "Accurate diagnosis of endoscopic mucosal healing in ulcerative colitis using deep learning and machine learning," *Journal of the Chinese Medical Association*, vol. 84, no. 7, pp. 678–681, 2021.

[12] J. Dhaliwal, L. Erdman, E. Drysdale et al., "Accurate classification of pediatric colonic inflammatory bowel disease subtype using a random forest machine learning classifier," *Journal of Pediatric Gastroenterology and Nutrition*, vol. 72, no. 2, pp. 262–269, 2021.

[13] E. Klang, Y. Barash, R. Y. Margalit et al., "Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy," *Gastrointestinal Endoscopy*, vol. 91, no. 3, pp. 606–613.e2, 2020.

[14] K. Takenaka, K. Ohtsuka, T. Fujii et al., "Development and validation of a deep neural network for accurate evaluation of endoscopic images from patients with ulcerative colitis," *Gastroenterology*, vol. 158, no. 8, pp. 2150–2157, 2020.

[15] J. P. S. Ferreira, M. de Mascarenhas Saraiva, J. P. L. Afonso et al., "Identification of ulcers and erosions by the novel Pillcam™ Crohn's capsule using a convolutional neural network: a multicentre pilot study," *Journal of Crohn's and Colitis*, vol. 16, no. 1, pp. 169–172, 2022.

[16] S. Soffer, A. Ben-Cohen, O. Shimon, M. M. Amitai, H. Greenspan, and E. Klang, "Convolutional neural networks for radiologic images: a radiologist's guide," *Radiology*, vol. 290, no. 3, pp. 590–606, 2019.

[17] Z. Ding, H. Shi, H. Zhang et al., "Gastroenterologist-level identification of small-bowel diseases and normal variants by capsule endoscopy using a deep-learning model," *Gastroenterology*, vol. 157, no. 4, pp. 1044–1054.e5, 2019.

[18] L. Peyrin-Biroulet, A. Bressenot, and W. Kampman, "Histologic remission: the ultimate therapeutic goal in ulcerative

colitis?," *Clinical Gastroenterology and Hepatology*, vol. 12, no. 6, pp. 929–934.e2, 2014.

[19] M. F. Neurath and S. P. Travis, "Mucosal healing in inflammatory bowel diseases: a systematic review," *Gut*, vol. 61, no. 11, pp. 1619–1635, 2012.

[20] R. W. Stidham, W. Liu, S. Bishu et al., "Performance of a deep learning model vs human reviewers in grading endoscopic disease severity of patients with ulcerative colitis," *JAMA Network Open*, vol. 2, no. 5, article e193963, 2019.

[21] R. T. Sutton, O. R. Zai Ane, R. Goebel, and D. C. Baumgart, "Artificial intelligence enabled automated diagnosis and grading of ulcerative colitis endoscopy images," *Scientific Reports*, vol. 12, no. 1, p. 2748, 2022.

[22] H. Yao, K. Najarian, J. Gryak et al., "Fully automated endoscopic disease activity assessment in ulcerative colitis," *Gastrointestinal Endoscopy*, vol. 93, no. 3, pp. 728–736.e1, 2021.

[23] K. Gottlieb, J. Requa, W. Karnes et al., "Central reading of ulcerative colitis clinical trial videos using neural networks," *Gastroenterology*, vol. 160, no. 3, pp. 710–719.e2, 2021.

[24] P. Bossuyt, H. Nakase, S. Vermeire et al., "Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density," *Gut*, vol. 69, no. 10, pp. 1778–1786, 2020.

[25] T. Aoki, A. Yamada, K. Aoyama et al., "Automatic detection of erosions and ulcerations in wireless capsule endoscopy images based on a deep convolutional neural network," *Gastrointestinal Endoscopy*, vol. 89, no. 2, pp. 357–363.e2, 2019.

[26] E. Klang, A. Grinman, S. Soffer et al., "Automated detection of Crohn's disease intestinal strictures on capsule endoscopy images using deep neural networks," *Journal of Crohn's and Colitis*, vol. 15, no. 5, pp. 749–756, 2021.

[27] Y. Barash, L. Azaria, S. Soffer et al., "Ulcer severity grading in video capsule images of patients with Crohn's disease: an ordinal neural network solution," *Gastrointestinal Endoscopy*, vol. 93, no. 1, pp. 187–192, 2021.

[28] M. Löwenberg and M. Van Der Vlugt, "Risk of progression of low-grade dysplasia to advanced neoplasia in inflammatory bowel disease," *Journal of Crohn's and Colitis*, vol. 13, no. 12, pp. 1483–1484, 2019.

[29] J. M. Huguet, L. Ferrer-Barceló, P. Suárez et al., "Colorectal cancer screening and surveillance in patients with inflammatory bowel disease in 2021," *World Journal of Gastroenterology*, vol. 28, no. 5, pp. 502–516, 2022.

[30] Y. Maeda, S. E. Kudo, N. Ogata et al., "Can artificial intelligence help to detect dysplasia in patients with ulcerative colitis?," *Endoscopy*, vol. 53, no. 7, pp. E273–e274, 2021.

[31] V. Villanacci, T. L. Parigi, R. Del Amor, P. Mesguer Esbrí, X. Gui, and A. Bazarova, "OP15 a new simplified histology artificial intelligence system for accurate assessment of remission in ulcerative colitis," *Journal of Crohn's and Colitis*, vol. 16, no. Supplement_1, pp. i015–i017, 2022.

[32] X. Gui, A. Bazarova, R. Del Amor, M. Vieth, G. de Hertogh, and V. Villanacci, "PICaSSO histologic remission index (PHRI) in ulcerative colitis: development of a novel simplified histological score for monitoring mucosal healing and predicting clinical outcomes and its applicability in an artificial intelligence system," *Gut*, vol. 71, no. 5, pp. 889–898, 2022.

[33] D. Sahoo, L. Swanson, I. M. Sayed et al., "Artificial intelligence guided discovery of a barrier-protective therapy in inflammatory bowel disease," *Nature Communications*, vol. 12, no. 1, p. 4246, 2021.

[34] F. Gul, S. N. H. Kazmi, K. Abbas, S. Saeed, and J. Basit, "The future of robotic surgery for inflammatory bowel diseases," *Annals of Medicine and Surgery*, vol. 81, article 104476, 2022.

[35] S. Hota, S. Parascandola, S. Smith, M. M. Tampo, R. Amdur, and V. Obias, "Robotic and laparoscopic surgical techniques in patients with Crohn's disease," *Surgical Endoscopy*, vol. 35, no. 8, pp. 4602–4608, 2021.

Hindawi
Gastroenterology Research and Practice
Volume 2022, Article ID 9285238, 9 pages
https://doi.org/10.1155/2022/9285238

Research Article

Performance of Deep Learning-Based Algorithm for Detection of Pediatric Intussusception on Abdominal Ultrasound Images

Zheming Li ,[1,2,3,4] Chunze Song,[3,5] Jian Huang ,[1,2,3] Jing Li ,[1,2,3] Shoujiang Huang,[3,6] Baoxin Qian,[7] Xing Chen,[8] Shasha Hu,[9] Ting Shu ,[10] and Gang Yu [1,2,3,4]

[1]*Department of Data and Information, The Children's Hospital Zhejiang University School of Medicine, Hangzhou 310052, China*
[2]*Sino-Finland Joint AI Laboratory for Child Health of Zhejiang Province, China*
[3]*National Clinical Research Center for Child Health, Hangzhou, China*
[4]*Polytechnic Institute, Zhejiang University, 866 Yuhangtang Rd, Hangzhou 310058, China*
[5]*Department of Ultrasound, The Children's Hospital Zhejiang University School of Medicine, Hangzhou 310052, China*
[6]*Department of Day Surgery, The Children's Hospital Zhejiang University School of Medicine, Hangzhou 310052, China*
[7]*Huiying Medical Technology (Beijing), Beijing 100192, China*
[8]*Hangzhou Normal University, 310052 Hangzhou, China*
[9]*The Children's Hospital Zhejiang University School of Medicine, Hangzhou 310052, China*
[10]*National Institute of Hospital Administration, NHC, Beijing 100044, China*

Correspondence should be addressed to Ting Shu; nctingting@126.com and Gang Yu; yugbme@zju.edu.cn

Received 30 June 2022; Revised 21 July 2022; Accepted 25 July 2022; Published 12 August 2022

Academic Editor: James H. Tabibian

Background and Aims. Diagnosing pediatric intussusception from ultrasound images can be a difficult task in many primary care hospitals that lack experienced radiologists. To address this challenge, this study developed an artificial intelligence- (AI-) based system for automatic detection of "concentric circles" signs on ultrasound images, thereby improving the efficiency and accuracy of pediatric intussusception diagnosis. *Methods*. A total of 440 cases (373 pediatric intussusception and 67 normal cases) were retrospectively collected from Children's Hospital affiliated to Zhejiang University School of Medicine from January 2020 to December 2020. An improved Faster RCNN deep learning framework was used to detect "concentric circle" signs. Finally, independent validation set was used to evaluate the performance of the developed AI tool. *Results*. The data of pediatric intussusception were divided into a training set and validation set according to the ratio of 8 : 2, with training set (298 pediatric intussusception) and validation set (75 pediatric intussusception and 67 normal cases). In the "concentric circle" detection model, the detection rate, recall, specificity, and $F1$ score assessed by the validation set were 92.8%, 95.0%, 92.2%, and 86.4%, respectively. Pediatric intussusception was classified by "concentric circle" signs, and the accuracy, recall, specificity, and $F1$ score were 93.0%, 92.0%, 94.1%, and 93.2% on the validation set, respectively. *Conclusion*. The model established in this paper can realize the automatic detection of "concentric circle" signs in the ultrasound images of abdominal intussusception in children; the AI tool can improve the diagnosis speed of pediatric intussusception. It is necessary to further develop an artificial intelligence system for real-time detection of "concentric circles" in ultrasound images for the judgment of children with intussusception.

1. Introduction

Intussusception is a kind of pediatric surgical acute abdomen, which is relatively common in clinic and mainly affects children under 2 years old [1–3]. Among the high number of pediatric emergency abdominal patients, the average annual incidence of intussusception in children in China is 418.1/100,000 [4, 5], which has exceeded the global average of 74/100,000 [6]. Intussusception mainly refers to the phenomenon of interlocking two intestinal tubes, early and timely diagnosis and active and correct treatment can prevent intestinal necrosis and relieve the pain of children [7, 8]. Ultrasound (US) as a noninvasive and painless examination method is easy to be accepted by children and their families [9, 10]. The ultrasound images of typical children intussusception can be summarized as two signs [11]: one is the "concentric circle" sign on the cross section, and the other is the "sleeve sign" sign on the vertical section. Doctors mostly judge whether patients have intussusception problem by identifying the "concentric circle" sign [12]. However, the increasing ultrasound image data also brings burden to doctors' diagnosis. In recent years, with the depth of the study in the field of deep learning applications, deep learning technology is often used in the field of fast and intelligent image processing, such as image classification and detection. Deep learning simulates human vision mechanism, and has the advantages of fast detection speed and low cost, especially in the field of medical imaging has made breakthrough progress. Many studies have shown that AI achieved surprising results. Qian et al. [13] developed an explainable deep-learning system based on multimodal breast-ultrasound images and predicted BI-RADS scores for breast cancer as accurately as experienced radiologists. Sui et al. [14] developed a deep-learning AI model (ThyNet) to differentiate between malignant tumours and benign thyroid nodules and aimed to investigate how ThyNet could help radiologists improve diagnostic performance and avoid unnecessary fine needle aspiration. Tiyarattanachai et al. [15] developed a deep learning network for the detection and diagnosis of focal liver lesions from ultrasound images, AI model detected and diagnosed common focal liver lesions. For diagnosis of hepatocellular carcinomas, the AI model yielded sensitivity, specificity, and negative predictive value of 73.6%, 97.8%, and 96.5% on the internal validation set. Although artificial intelligence is widely used in the detection and classification of lesions in breast, thyroid, liver and other ultrasound images, the application of artificial intelligence in pediatrics is still in its infancy.

Region Convolutional Neural Network features (RCNN) with convolutional neural network features were proposed for target detection by Girshick et al. [16] in 2014. Since then, target detection has started to evolve at an unprecedented rate. Although RCNN has made great progress, it requires a large amount of redundant feature computation, resulting in extremely slow detection. In 2015, Girshick proposed the Fast RCNN detector [17], which is a further improvement to RCNN. Fast RCNN allows us to train both the detector and the bounding box regressor in the same network configuration. The detection speed is more than 200 times faster than that of RCNN. Although Fast-RCNN successfully computes the feature mapping only once for the whole image, its detection speed is still limited by region proposal network. In 2015, Ren et al. proposed the Faster RCNN detector [18], which is the first end-to-end and the first near real-time deep learning detector. Since then Faster RCNNs have been widely used for detection tasks.

Inspired by the classical detection network Faster RCNN, this paper developed a model more suitable for detecting "concentric circles" in ultrasound images. To the best of our knowledge, this is the first attempt that makes use of deep learning to diagnose intussusception in children, it makes up for the blank of using deep learning for diagnosing pediatric intussusception based on ultrasound images.

2. Material and Method

The retrospective study was approved by institutional ethics committee of Children's Hospital Affiliated to Zhejiang University School, and a waiver for informed consent was provided.

2.1. Patient Selection

2.1.1. Selection Criteria

(1) Patients diagnosed with intussusception and normal cases

(2) The abdominal ultrasound image data were complete

(3) Ultrasound images are clear

2.1.2. Exclusion Criteria

(1) Previous abdominal surgery

(2) There are motion artifacts and foreign bodies in the ultrasound image

This study retrospectively collected the images and clinical data of pediatric intussusception in admitted to Children's Hospital Affiliated to Zhejiang University School of Medicine from January 2020 to December 2020. A total of 440 children were included in this study, with 372 pediatric intussusception and 67 normal children.

2.2. US Image Acquisition. With ultrasound examination instruments and equipment examination methods: Philip IE33 and iuEilite color ultrasound instrument used linear array probe frequency L 5~12 and convex array probe frequency 1~5 MHz, the child was in supine position, with a linear array probe combined with a convex array probe for detailed scan of the entire abdominal bowel. After the lesions were found, multisection scanning, such as longitudinal and transverse resection, was performed to observe the lesions in real time with the change of body position of the children. After the lesions were clearly displayed, the location of the mass was recorded, the diameter of the "concentric circle " sign and the length of the "sleeve sign" were measured, and the mesenteric lymph nodes in the lesions were observed.

The echogram of pediatric intussusception showed "concentric circle" signs. Signs of "concentric circles" at different angles are shown in Figure 1. For naked eye observation, some "concentric circle" signs are obvious and easy to observe, such as Figure 1(a). However, the image comparison between the target area and background area is too low and the edge is blurred, and the "concentric circle" signs are irregular, such as Figure 1(b); these characteristics are also difficult to be recognized by human experts, so the detection of "concentric circle" signs is a challenging task.

2.3. Data Processing. All data were static images taken by the subjects during ultrasound, and the images were stored in DICOM format. The dataset consists of images from different scales, with an average image size of 1341×864 pixels. Because of the subjectivity of probe strafing, some of these images contain "concentric circle" signs, while others do not.

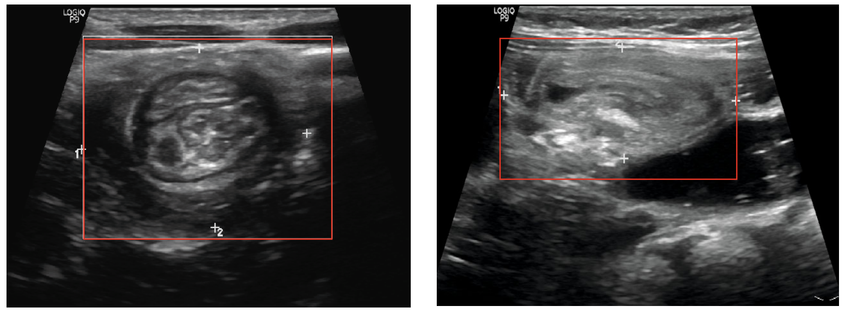

(a) (b)

Figure 1: Signs of “concentric circles” at different angles. A child diagnosed with intussusception (female, 2 years old) presents with abdominal pain, vomiting, hematochezia, and abdominal mass. The “concentric circle” area is marked with a white cross: (a) easy to observe and (b) hard to observe.

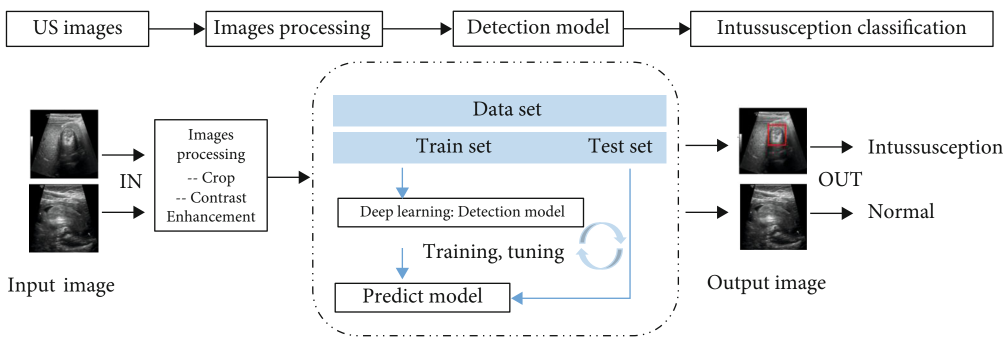


Figure 2: Overall flow chart.

In the process of image preprocessing, all the subject identification information and peripheral regions in the ultrasound images are cropped to ensure that the cropped image only contains the fan-shaped ultrasound region.

2.4. AI Model Development. This study is divided into two parts: the first part includes image preprocessing and detection of “concentric circle” signs using an improved Faster RCNN network; the second part is based on the detection of “concentric circle” signs to complete the classification of pediatric intussusception and normal cases. The overall flow chart of this study is shown in Figure 2. The original DICOM images were first cropped and contrast enhanced; then, the images after processing were fed into the detection model for training, and based on the model’s prediction of the “concentric circle” signs, the final category and its probabilities were predicted to give a higher category probability and a prediction score for patients with intussusception.

2.5. Design of the Detection Model. Inspired by the classical detection network Faster RCNN, this paper developed a model more suitable for detecting “concentric circle” signs in ultrasound images. In order to distinguish standard “concentric circles” and nonstandard “concentric circles” better, this paper adds a jump connection to the convolutional neural network. By combining the shallow and deep features of the convolutional neural network with the jumping connection layer, the detector can detect the regions with insignificant background differences and no obvious “concentric circles” signs. Based on the Faster RCNN network model, this paper uses layer connection to connect the features collected at conv3 and conv5 layers of VGG16 [19] (network structure is shown in Figure 3), so as to fuse shallow information and deep features to better mine semantic features in images. We changed the last fully connected layer to predict two categories: the “concentric circle” area and the background. In addition, we retrain the last fully connected layer and calculate the coordinates and confidence of the “concentric circle” region and background.

2.6. Experiments. The model was trained by the method of supervision training. To obtain the corresponding image label, the position of the “concentric circle” signs in the ultrasound image was traced by an experienced sonographer and verified by another expert label to ensure accuracy. We used PyTorch framework to train the model for 120 iterations (120 epochs, each with 1000 iterations) on 2 GeForce RTX 1080 Ti GPUs. For model training, the Batch size was set to 64, and the initial learning rate was set to 0.01. Use warm restart learning rate to adjust learning rate [20]. Cosine function can be used to reduce learning rate. In the cosine function, with the increase in x, the cosine value decreases slowly at first, then accelerates, and slowly decreases again. This kind of decline mode can be combined with the learning rate to produce good results in a very effective way of calculation. The training time was 1 D 0 H 24 min. The loss curve of model training is shown in

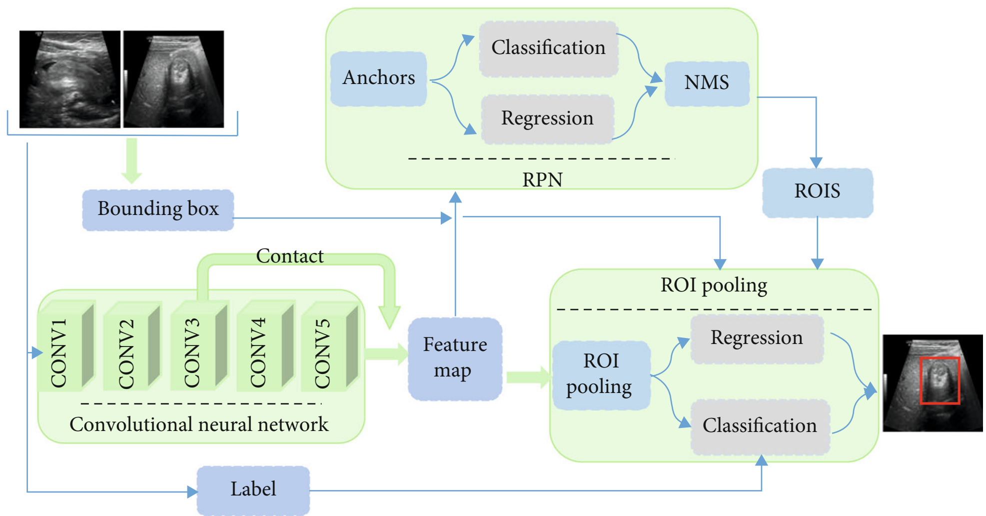


Figure 3: Network structure. The main structure of the network consists of three stages: backbone for feature extraction network, region proposal network (RPN), and region of interest (ROI) pooling. Input the image into the VGG16 to get the feature map, use RPN to generate the anchors, after Nonmaximum Suppression (NMS) to obtain ROIs, project the ROIs onto the feature map to get the feature matrix, scale each feature matrix to 7×7 through the ROI pooling, and then flatten the feature map to get the prediction result through a series of fully connected layers.

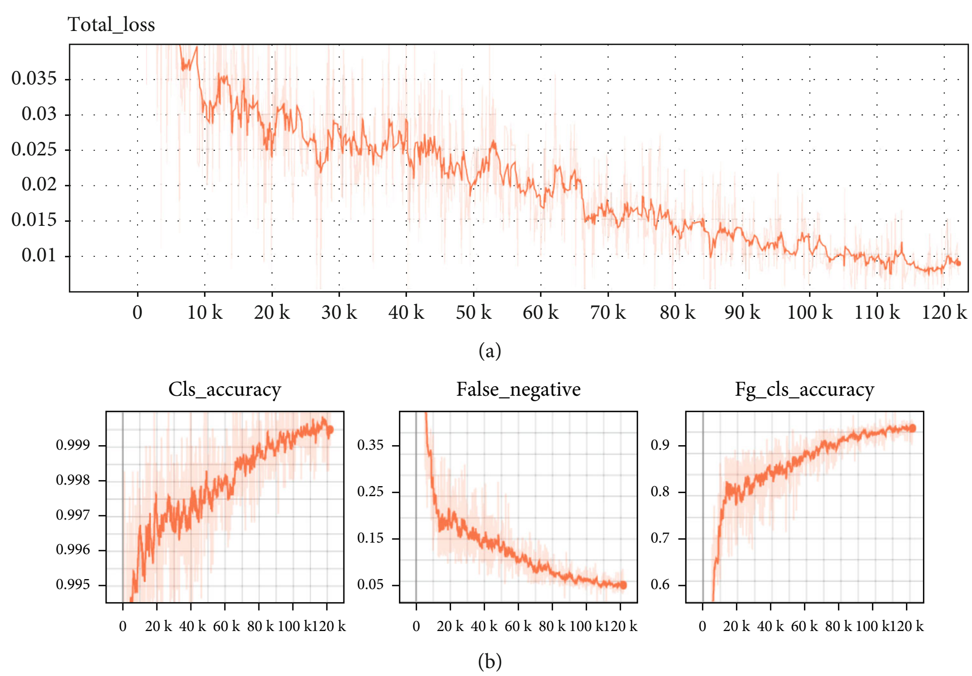


Figure 4: Training details. (a) Loss curves of training sets under different iterations and (b) variation curves of different evaluation index values of training sets under different epochs.

Figure 4(a). The loss curve shows that when the number of iterations is 120,000, the model is in the fitting state. On the other hand, the model is in an ideal training state when 120,000 iterations are reached. The model with 120,000 training times was selected as the final detector for testing. Figure 4(b) shows that with the increase in epoch times, each evaluation index tends to be stable.

2.7. Statistical Analysis. SPSS 22.0 was used to identify differences in clinical features between patients with intussusception and the normal case group. Continuous variables were expressed as mean ± standard deviation, and the two-sided Student t-test was used to compare whether there were significant differences between groups. Discrete variables were expressed as counts (percentages), and Pearson's chi-

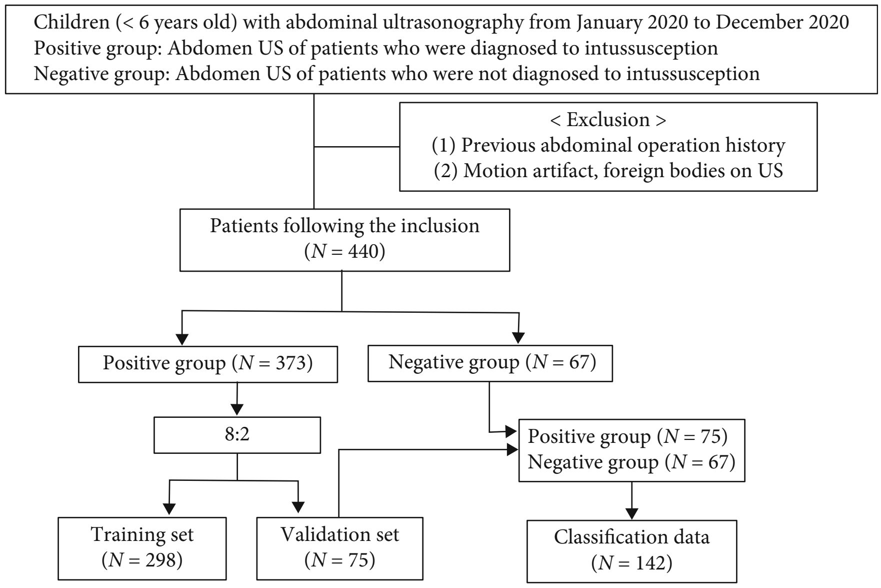


Figure 5: Flow chart of data collection.

square test was used. If $P < 0.05$, the variables were considered to be significantly different among different groups.

3. Results

3.1. Clinical Characteristics. According to the criteria, 440 patients (306 males and 134 females, mean age 33.9 ± 29.1 months) were enrolled. The flow chart for data collection is shown in Figure 5.

A total of 440 patients were included in this study, including 373 patients with intussusception (male : female = 266 : 107, mean age: 28.6 ± 21.5 months) and 67 patients in the normal group (male : female = 40 : 27, mean age: 39.1 ± 36.6 months). Table 1 shows the demographics of the two groups. Univariate analysis of age and sex by independent sample t-test and Pearson's chi-square test showed that the patients with intussusception were younger than normal in demographic characteristics ($P = 0.001$). The majority of intussusception patients were boys ($P < 0.001$), and the results were statistically significant.

3.2. Data Split. A total of 440 cases were included in this study, with 373 pediatric intussusception patients and 67 children. There were 5 to 10 ultrasound images in each case, no ultrasound images with "concentric circle" signs in the normal group, and 1 to 3 ultrasound images with "concentric circle" signs in each intussusception patient, and a total of 715 ultrasound images with "concentric circle" signs were labeled. The data set was divided by 8 : 2, with 80% of the data as a training set (2325 images of 298 pediatric intussusception patients, including 575 positive samples); 20% of the data (586 images of 75 pediatric intussusception patients, including 140 positive samples) were used as a validation set to evaluate the performance of the concentric circle signature detection model. 75 pediatric intussusception patients and 67 normal cases were classified to evaluate the generalization ability of the model

Table 1: Patient baseline data.

	Intussusception	Normal	P value
All ($N = 440$)			
Patient, n	373	67	
Age (months), mean [SD]	28.6 [21.5]	39.1 [36.6]	0.001
Sex, male, n (%)	266 (71.3)	40 (59.7)	<0.001

(for detection data, see Table 2, and for classification data, see Table 3).

3.3. Performance and Evaluation of the Detection Model. The detection rate of the "concentric circle" signs in each image was evaluated: if the prediction model generated a boundary box around the image and the box overlapped with the real location of the "concentric circle" signs, the "concentric circles" were judged to be correctly detected. In this study, different confidence levels were used to evaluate the detection effect. The detection effect of concentric circles varies with different confidence levels, as shown in Table 4. Accuracy (Acc) was calculated by dividing the true positive (TP) number of correctly detected concentric circle signs by the total number of concentric circle signs. When the model outputs a bounding box in an area that does not contain a "concentric circle" sign, the count is false positive (FP). Evaluation indicators are defined as follows:

$$
\begin{aligned}
\text{Acc} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \\
\text{Spe} &= \frac{\text{TN}}{\text{TN} + \text{FP}}, \\
\text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\
\text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\
F1\ \text{score} &= 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}.
\end{aligned} \tag{1}
$$

Table 2: Division of detection data.

Detection data		Total case	Total images	Positive images	Negative images
Train set	Intussusception	298	2325	575	1750
Validation set	Intussusception	75	586	140	446
Total number		373	2911	715	2196

Table 3: Division of classification data.

	Intussusception	Normal	Total case
Classification data	75	67	142

Table 4: Comparison of evaluation index results under different confidence thresholds.

Threshold	TP	TN	FP	FN	Acc	Spe	Recall	Precision
0.0	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.05	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.10	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.15	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.20	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.25	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.30	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.35	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.40	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.45	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.50	133	411	35	7	92.8%	92.2%	95.0%	79.2%
0.55	132	411	35	8	92.7%	92.2%	94.3%	79.0%
0.60	131	411	35	9	92.5%	92.2%	93.6%	78.9%
0.65	131	412	34	9	92.7%	92.4%	93.6%	79.4%
0.70	130	412	34	10	92.5%	92.4%	92.9%	79.3%
0.75	130	412	34	10	92.5%	92.4%	92.9%	79.3%
0.80	128	413	33	12	92.3%	92.6%	91.4%	79.5%
0.85	127	413	33	13	92.1%	92.6%	90.7%	79.4%
0.90	124	415	31	16	92.0%	93.1%	88.6%	80.0%
0.95	122	417	29	18	92.0%	93.5%	87.1%	80.8%

Finally, 0.5 was used as the cut-off point of confidence threshold. In the "concentric circle" sign detection model, the detection rate, recall rate, specificity, and $F1$ score evaluated by the validation set were 92.8%, 95.0%, 92.2% and 86.4%, respectively. Ultrasound images of pediatric intussusception in the validation data are shown in Figure 6.

ROC curve of the final "concentric circles" sign detection model on the validation set is shown in Figure 7. The closer the ROC curve is to the upper left, the better the classifier performs. The Area under the Curve (AUC) of the "concentric circle" detection results on the validation set in this paper was 95.1%. After obtaining the ROC curve of the final "concentric circles" sign detection model, we determined an optimal threshold of 0.821 according to the maximized Youden index (sensitivity + specificity − 1) from the validation set.

3.4. Performance and Evaluation of Classification. Based on the results of the concentric circle detection, the algorithm used input images one by one to predict whether there was a dichotomy between normal cases and pediatric intussusception. The model gives the predicted pediatric intussusception score by multiplying the probability of an object's existence in ROI by the probability of category, and the class with higher probability is predicted as the final class and its probability. The validation set consisted of 75 pediatric intussusception and 67 normal cases. The confusion matrices on the validation set is presented in Figure 8(a). Where the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) diagnoses were 69, 63, 6, and 4, respectively. Table 5 shows the comparison of classification performance on the validation set using different evaluation indexes. The AUC, accuracy, recall, specificity, and $F1$ score were 98.6%, 93.0%, 92.0%, 94.1%, and 93.2%, respectively, on the validation set. ROC curve of classification results is shown in Figure 8(b). The AUC of pediatric intussusception diagnosis on the validation set was 98.6%. The results showed that the dichotomy of normal cases and pediatric intussusception had good classification performance according to the detection results of "concentric circle" signs.

4. Discussion

In this paper, an improved Faster RCNN was applied in the detection of "concentric circle" signs in abdominal ultrasound for the diagnosis of pediatric intussusception. The visualization results show that the model effectively learns the signs of concentric circles. The detection rate, recall, specificity, and $F1$ score of the algorithm evaluated in the validation set were 92.8%, 95.0%, 92.2%, and 86.4%, respectively. In our study, the algorithm used input images one by one to predict whether "concentric circle" signs exist to classify patients as normal cases and pediatric intussusception. The model through an object exists in the ROI probability and categories probability multiplication to give prediction score for pediatric intussusception. The class with the highest probability is predicted to be the final class and its probability. The accuracy, recall, specificity, and $F1$ score of the concentric circle signs in the diagnosis of pediatric intussusception were 93.0%, 92.0%, 94.1%, and 93.2% in the internal test set, respectively.

There are some differences between our study and previous studies. Kim et al. [21] developed and tested the performance of a deep learning-based algorithm to detect ileocolic

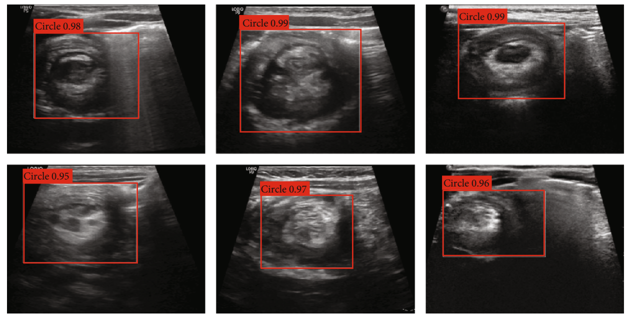


FIGURE 6: Ultrasound images of pediatric intussusception in the validation set.

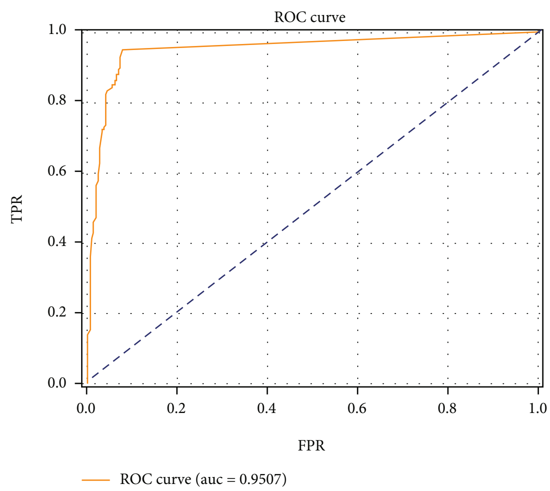


FIGURE 7: ROC curve of detection results on the validation set.

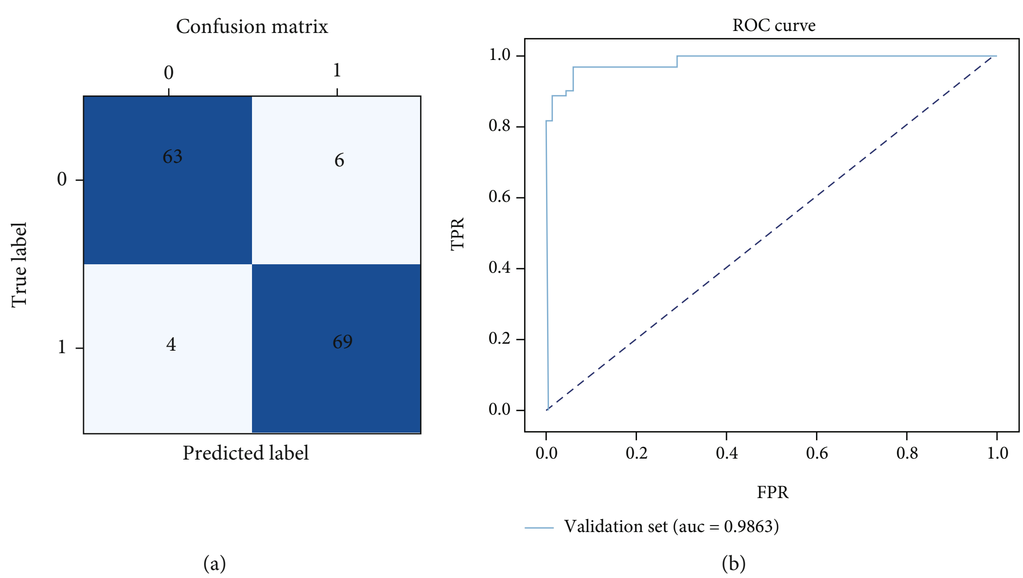


FIGURE 8: Classification results on the validation set. (a) Confusion matrix and (b) ROC curve.

intussusception using abdominal radiographs of young children, and a YOLOv3-based algorithm was developed to recognize the rectangular area of the right abdomen and to diagnose intussusception. The sensitivity of the algorithm was higher compared with that of the radiologists (0.76 vs. 0.46, $P = 0.013$). Compared with Kim et al.'s work, we used abdominal ultrasound images of children and the algorithm of Faster RCNN based on the detection of concentric circle signs to diagnose intussusception in children; the sensitivity of the algorithm is higher than that of Kim et al..

Table 5: Classification performance on the validation set using different evaluation indexes.

	AUC	ACC	Recall	Spe	F1 score	Youden index
Validation set	98.6%	93.0%	92.0%	94.1%	93.2%	0.813

Our study also had certain limitations: (1) in this study of deep learning for the diagnosis of intussusception in children on abdominal ultrasound images, the amount of data used by our algorithm is limited. Future work will focus on improving the result of deep learning for the detection of intussusception by adding more training data. (2) The data for this paper came from only one hospital. In the future, data from multiple medical institutions can be considered for external validation to verify the generalization performance of the model. (3) The interpretability of the deep learning model has always been criticized. In the follow-up work, some interpretable features will be added to improve the interpretability of the model.

5. Conclusion

In summary, this paper has developed a deep learning framework for the detection of "concentric circle" signs on ultrasound images. The model has good accuracy and reliability for "concentric circle" detection; in addition, it has a high accuracy in classifying pediatric intussusception based on the results of the detection, and a chart review was conducted to confirm that the imaging correctly identified the diagnosis.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Ethical Approval

This study was approved by the institutional review committee of participating hospitals, which exempted the requirement of informed consent of each patient.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contributions

Zheming Li and Chunze Song contributed equally to this study.

Acknowledgments

This work was partially supported by the National Key R&D Program of China (grant number 2019YFE0126200) and National Natural Science Foundation of China (grant number 62076218).

References

[1] National Health Commission of the People's Republic of China, "The reform and development of children's medical and health services (Guoweiyifa (2016) No. 21) [EB/OL]," 2016.

[2] National Health Commission of the People's Republic of China, *National Children'sMedicalCenter and National Children's Regional Medical Center Setup Plan (Guoweiban Yifa [2016] No. 31) [EB/OL]*, 2016.

[3] National Bureau of Statistics of China, *Demographic survey sample data [EB/OL]*, 2020.

[4] N. Liu, C. Yen, T. Huang et al., "Incidence and epidemiology of intussusception among children under 2 years of age in Chenzhou and Kaifeng, China, 2009-2013," *Vaccine*, vol. 36, no. 51, pp. 7862–7867, 2018.

[5] X. Z. Li, H. Wang, J. Song, Y. Liu, Y. Q. Lin, and Z. X. Sun, "Ultrasonographic diagnosis of intussusception in children: a systematic review and meta-analysis," *Journal of Ultrasound in Medicine*, vol. 40, no. 6, pp. 1077–1084, 2021.

[6] M. N. Patsikas, L. G. Papazoglou, and G. K. Paraskevas, "Current views in the diagnosis and treatment of intestinal intussusception," *Topics in Companion Animal Medicine*, vol. 37, article ???, 2019.

[7] X. Y. Zhang, Y. Gao, C. P. Li et al., "Observed and projected trends in paediatric health resources and services in China between 2003 and 2030: a time-series study," *BMJ Open*, vol. 7, no. 6, article e015000, 2017.

[8] L. I. Kelley-Quon, L. G. Arthur, R. F. Williams et al., "Management of intussusception in children: a systematic review," *Journal of Pediatric Surgery*, vol. 56, no. 3, pp. 587–596, 2021.

[9] W. L. Guo, S. F. Zhang, J. E. Li, and J. Wang, "Association of meteorological factors with pediatric intussusception in subtropical China: a 5-year analysis," *PLoS One*, vol. 9, no. 2, article e90521, 2014.

[10] E. A. Edwards, N. Pigg, J. Courtier, M. A. Zapala, J. D. MacKenzie, and A. S. Phelps, "Intussusception: past, present and future," *Pediatric Radiology*, vol. 47, no. 9, pp. 1101–1108, 2017.

[11] P. J. Strouse and A. L. Hryhorczuk, "Validation of US as a first-line diagnostic test for assessment of pediatric ileocolic intussusception: reply to letter from Cohen," *Pediatric Radiology*, vol. 40, no. 5, pp. 789-790, 2010.

[12] S. Henrikson, C. E. Blane, K. Koujok, P. J. Strouse, M. A. DiPietro, and M. M. Goodsitt, "The effect of screening sonography on the positive rate of enemas for intussusception," *Pediatric Radiology*, vol. 33, no. 3, p. 190, 2003.

[13] X. Qian, J. Pei, H. Zheng et al., "Prospective assessment of breast cancer risk from multimodal multiview ultrasound images via clinically applicable deep learning," *Biomedical Engineering*, vol. 5, no. 6, pp. 522–532, 2021.

[14] S. Peng, Y. Liu, W. Lv et al., "Deep learning-based artificial intelligence model to assist thyroid nodule diagnosis and management: a multicentre diagnostic study," *The Lancet Digital Heath.*, vol. 3, no. 4, pp. e250–e259, 2021.

[15] T. Tiyarattanachai, T. Apiparakoon, S. Marukatat et al., "Development and validation of artificial intelligence to detect and diagnose liver lesions from ultrasound images," *PLoS One*, vol. 16, no. 6, p. e0252882, 2021.

[16] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Te IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[17] R. Girshick, "Fast R-CNN," *IEEE International Conference on Computer Vision*, pp. 1440–1448, 2015.

[18] S. Ren, K. He, R. Girshick, and J. Sun, *Faster R-CNN: towards real-time object detection with region proposal networks*, IEEE Transactions on Pattern Analysis & Mach. Intell, 2016.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition [EB/OL].(2015-04-10)[2020-07-25]," 2014, https://arxiv.org/abs/1409.1556.

[20] H. Li, M. Nijiati, R. Zhou et al., *Active/secondary pulmonary tuberculosis classification model based on deep learning of CT*, Images and Its Clinical Applications, 2022.

[21] S. Kim, H. Yoon, and M.-J. Lee, "Performance of deep learningbased algorithm for detection of ileocolic intussusception on abdominal radiographs of young children," *Scientific Reports*, vol. 9, pp. 1–8, 2018.